UNITED
NATIONS

E

Economic and Social Council

Distr.
LIMITED

E/CONF.85/L.70
22 July 1992

ENGLISH ONLY

SIXTH UNITED NATIONS CONFERENCE
ON THE STANDARDIZATION OF
GEOGRAPHICAL NAMES
New York, 25 August-3 September 1992
Item 7 of the provisional agenda*

TERMINOLOGY IN THE STANDARDIZATION OF GEOGRAPHICAL NAMES

Remarks on definitions

Paper submitted by Hungary**

* E/CONF.85/1/Rev.1

** Prepared by Ervin Földi, Hungary.

92-33424 3539e (E)

/...

The following *acronyms* are used in the remarks:

TECHTER: Technical terminology employed in the standardizationof geographical names. UN Secretariat, Department Conference Services, Translation Division, Glossary No. 330/Rev.2, 24 July 1987;

TOPTER: Glossary of toponymic terminology. Edited by Naftali Kadmon. UN Group of Experts on Geographical Names, Working Group on Terminology. May, 1991.

Please, disregard any remarks that are no more valid if a change had been made since in a new version of TOPTER!

1. The importance of having clearly defined terms in the work of national and international standardization is unquestionable. It seems also reasonable that TECHTER or any material on terms of a constantly developing activity should be revised from time to time as our knowledge increases and circumstances change.

However, on the basis of TOPTER I have the impression that a number of unnecessary changes will be submitted to this Conference, while some important aspects will not be reflected in the terms included as well as in their definition. I am aware of the risk, commenting materials and details that may have been superseded, but I hope, my remarks could be of use even so.

2. The necessity of a close contact with computers for everybody engaged with geographic names is an imperative of our age. However, I consider it as an evidence, that a glossary on toponymic terminology should not contain *elementary terms of the use of computers*. The following terms in TOPTER fall into this category:

002 address
017 alphanumeric
021 batch processing
064 default value

102 file, computer
106 format, computer
134 interactive
135 interface
181 menu, computer
213 noise, graphic
238 raster mode
241 record, computer
267 software
319 vector mode

3. Remarks on individual terms and on their definition

allonym

As a consequence of this and other definitions, names of a feature located in a region with a *non-official principal language* are both *allonyms* and *endonyms*, moreover, the name in the *non-official principal language* falls into a third category as well, it is also an *exonym*, on the basis of the definitions presented in this Glossary. See also remarks on *endonym; endonym, standardized*. This is a situation that has to be avoided.

allonym, standardized

The remarks given on *allonym* are also valid here in connection with the term *standardized endonym*. See also remarks there.

alphabet, transliteration

The expression "sometimes limited in extent" is meaningless, because we have no information, to what basis the limitaion occurs. If an *alphabet* can serve *transcription* or *transliteration* without limitation, it might not called limited. As there is no Arabic /p/, there is no need for it in the *transliteration alphabet* used in the *romnization* of Arabic, therefore, it is complete even without the P. It is incomplete only in terms of any Roman alphabet containing P, but this fact is of no significance here.

alphabet, transcription

The expression "sometimes limited in extent" is meaningless, see remarks on *transliteration alphabet*.

category, topographic; class, feature

If *feature classes* form *topographic categories*, then why is defined *topographic category* with the same meaning? Definition should state that a *topographic category* consists of *feature classes*. The example "stream" appearing in both definitions, also

suggests, that the two terms are identical. I consider both terms unnecessary as they coincide with *generic term*.

character, simplified

As to Chinese, *simplified character* has (also) another meaning: A newly designed character having the same meaning and pronounciation, but considerably fewer strokes, than the original one, eg. 东 for 東 , 国 for 國 , 归 for 歸 , etc. These pairs of characters are at the same time *variant characters*.

character, vowel

There should be "syllabic" instead of "non-alphabetic writing system" in the definition. Cf. *writing system*

coded representation

It should be mentioned that the coding may be alphanumeric, and graphic, the latter one used sometimes in indexes to give *feature categories*. Alphanumeric codes are the ISO country codes (two letter, three letter, and three numerals).

conversion, names

The addition of "names" seems unnecessary, as it is difficult to apply in most of the definitions referring to this term, and it is already left in a considerable number. Cf.

This term was originally included to solve the contradiction, that in the case of *logograms* there is no *script* from which anyone could *transform* directly into another *script* (cf. *logogram*), in contrary to the other two *writing systems*. It can be seen in the definition for *conversion* in TECHTER, that it mentions "phonological and/or morphological elements of a language or the graphic symbols of another writing system..." but it does not refer to a script! This definiton of the TECHTER is more practical in another aspect, too: it gives a possibility to include situations, when an *alphabetic* or *syllabic script* makes direct *transliteration* or *transcription* impossible, eg. the *script* contains no information in Devanagari on the use of the inherent vowel, the same is true for Amharic, besides, there is no information on double consonants, not to speak about the defective alphabetic scripts. All these examples show that *conversion* as defined in TECHTER covered situations that have been left out of consideration TOPTER. -- In other words: the definition in TOPTER is incorrect.

coordinates, topographic

I wonder why the term *grid coordinates* cannot be used. Even the term *rectangular coordinates* would be much better.

Besides, under *coordinates* we understand in cartography usually *coordinate values*, and not the *grid* itself, as this definition says (cf. Term No. 321.17 in Multilingual Dictionary of Technical Terms in Cartography, Wiesbaden, 1973).

The necessity of improving this definition is shown from the definition for UTM grid, where the term *topographic coordinates* is forgotten. It suggests, that the UTM grid is not formed by *topographic coordinates*, but by something else. Reference to each other is also missing.

Another question: why *geographical coordinates* (in the aforementioned Dictionary: *graticule co-ordinates*) are not included in this Glossary? Just because in many cases they are exclusively used in locating *geographic names*, while *grid coordinates* usually appear together with *geographical coordinates* only?

data element, data item

I cannot see any difference between the actual meaning of the two terms. According to the definition. the item is the content of the element, however, the example is in both cases a date, and it seems quite clear, that both data occupie a complete data field. They must be synonyms.

data portability

"The capacity of running the same program ..."

diacritic

It seems that TOPTER definition in contrary to TECHTER excludes *diacritics* (at least in some languages) which give the position of a stress without a *phonemic* value. If this is not true, examples of this kind should also be given.

element, generic

This definition is incorrect in TOPTER because it has been incorrect already in TECHTER. This is clearly shown if we take the example Port-au-Prince (correctly with hyphens), because Port in this name is not a *generic term*, as the feature is the capital of Haiti, and the port is only a smaller part of the city, therefore the *generic term* would be "populated place". The second part of the definition in TOPTER is a correct statement, and the whole definition could be improved by not referring to *generic term*.

element, specific

Here again, reference to *generic term* should be left, as many geographic names do not contain a *generic term* in the sense defined in this Glossary.

endonym

It seems that the definition of this term is unnecessarily broad. Let me take a theoretical example. Somewhere in East Central Europe there is a region, outside of Hungary, with the Hungarian as the *non-official principal language*. The Hungarian name of a commune (village) or of a city in this region, is an *endonym* (according to the definition here), and, at the same time, an *exonym* (according to the definition of exonym). Of course, there is another name present, the *official name*, i.e. the name in the *official languge* of this region.

In other words: if the term *exonym* is strictly connected with *standardization*, this fact cannot be left from the definition of the *endonym*. Otherwise it is impossible to seperate them.

The difficulties described here originate from the fact, that in this Glossary the use of the term *minority language* is deliberately avoided and substituted by the terms *non-official language*, and *national language* (see also remarks there).

endonym, standardized

This term also seems to be the result of a misunderstanding. Standardization is needed when *allonyms* are present. An *endonym*, as defined above, is a single name with the exception of the presence of more than one *language*. The example Kingston-upon-Hull and Hull shows that they are both English language *allonyms*. Contrary to the conception of this definition, a *standardized endonym* leads to the situtaion described in the definition of the term *standardized allonym*.

exonym

The definition has been slightly modified in TOPTER as compared to TECHTER. Excluding variants in diacritics only, and converted forms is logical. On the other hand, joining *conversion*, (in terms of TOPTER *transfform* should be used!) some dangers occur.

The definition suggests to consider all *conversions* (or *transformations*) are exceptions, while in the practice, deviation form any *systematic conversion* may produce exonyms. Example: Both *Moscou* and *Moskva* are conversions, but the first one is also an exonym in terms of *standardization*. This means, that another term is in the background: Conversion exonym: A name converted into a specific target script by using a different *transcription* or *translite-transliteration key* as compared to names in the same context.

Another problem is that the possibility using *minority language* toponyms parallel to the name in the *official language* produces a situation not reflected in this definition (cf. Report

by Hungary, Annex: Legal background of name standardization). Definiton of TOPTER could be corrected simply by writing "standardized name or names" instead of "standardized name".

Relating to other problems, see remarks on *allonym*, *standardized allonym*, *endonym*, *standardized allonym*.

exonymization

This definition leads to a dangerous and delicate territory. The *history* of a *toponym* has been included.

The definition suggests that an *exonym* originates from the need to substitute an *endonym* (at this point I do not repeat why the TOPTER definition for *endonym* is wrong, see there). This is clearly shown looking at the definition of *names transformation*, where it is stated that *exonymization* is a similar act as *translation*, *transcription* or *transliteration*. This is not true, as there are many examples in the history of *geographical names* considered at present as *exonyms*, when *endonyms* had to be created for topographic features in case of newly acquired territories, i.e. *endonyms* were substituted for *exonyms*.

Another question: where does this substitution happen? In the language, or on the map? If it is intended to refer to a map, it must be clearly stated and it has to be decided whether the use of an *exonym* below the *endonym* on the map in a smaller size, and perhaps in brackets, is *exonymization* or not (at present, it is not, because no substitution happens).

feature, hydrographic

In TOPTER this term was limited for the oceans and seas, and the term *hydrological feature* has been included for inland waters. In TECHTER both category appear under one term, and the examples suggest, that *marine features* are excluded, though the definition includes all features consisting of water.

font

Computer printers use *font* meaning style, and the term *pitch* meaning sizes. Cf. *typeface*.

hyphenization

Meaning (a) is the same as that of *syllabification*, therefore it should be dropped.

Second part of the definition (b) is unnecessary and misleading. "Preventing change of word order" may be of use in some languages, but in other languages, such as German or Hungarian, the use of hyphens is a question syntax. In Hungarian geographical names the *hyphen* shows in most cases that the

elements joined together are similar to elements of a composite word, but actually not, as elements of a composite word are written together in *one word*. The definition is misleading, because it gives the impression that the only function of a *hyphen* could be prevention.

international names rendition

A map described here should use the *ideographic script* for Chinese, and the *syllabic script* for India etc. This is a nonsense! The definition for *multilingual lettering* in TECHTER is much better, even if the word *multilingual* is not the best one. I would prefer to keep words in terms that have been used in the English language for a considerable time, even if the word could be substituted by a more precise one but never used before. This is the case with many terms in TOPTERM, eg. *names rendition*, *multiscriptual*, *topographic coordinates*, *allonym* etc.

IPA

IPA can serve phonemic transcription as well.

language, source; language, target

According to the definition of *names transformation*, this definiton should be put "...a name is transformed to ..." and "...a toponym may be ... transformed..." resp.

letter, basic

In the example "Polish l as against ł" should stand.

lettering, map

I consider it as an unnecessary precision to make distinction between *writing systems* in such a purely technical matter. even if the term *letter* belongs to *alphabetic writing system* only. See remark on *international map rendition*.

linguistics

Syntax being a term in TOPTERM should be mentioned here.

map script

As far as I know, the word *script* in the English cannot describe an activity, it is not a verb like *lettering*. I dont think that in the interest of extremely precise scientific categorization a new type of English language should be invented. Either the term should be changed to *name placing* (as it is written in the definition) or the definition should be changed to "The totality of toponyms and other ...on maps"

Consideration should be given to the definition of *script* in this Glossary as well.

map script, multiscriptual

This method is much more frequent in gazetteers, therefore, the use of the word "map" in the term is less justified, than to put it the other way: *multiscriptual gazetteer* (term), "...in a gazetter or (sometimes) on maps .." (definition).

marker

As mentioned in connection with the term *radical*, a *logogram* cannot have a *marker* "...used only in combination ...". See remarks on *radical*.

name, composite

Example should be given, when the free *morphemes* are in one word, eg. Kemijoki (Finnish 'Kemi River').

phoneme

I wonder why the word "functionally" is put into brackets when this term gives the essence of the entire notion as opposite to non-functional *allophones*.

phonology

Definition is entirely incorrect. *Phonology* deals with the semiotic aspect of sounds as distinct from *phonetics*. *Phonology* aims at studying not only the speech sound, but its function in the speech structure. The present definition is a variation on the theme *phonetics*.

radical, radix

The *radical* or *radix* is in Chinese one of the two equal parts of most *logograms*. Therefore, the example, and the definition is not right: the *radical* "tree" is combined with other *logograms*, which have an independent meaning, too. Eg. the *logogram* "tree" is combined with the *logogram* "orchid", and the result is the *logogram* for "magnolia"; or the *logogram* "tree" is duplicated into one *logogram*, and the new *logogram* means "forest". There is also a *logogram* consisting of three original "tree" *logograms*, meaning "dense forest, shadow, etc." In other words: if we accept the definition *marker* ("used only in combination"), then the definition for *radical* is incorrect.

romanization

In terms of the present Glossary, the correct definition would be: "*Transformation* from a non-Roman script ..."

script

The TOPTER definition is actually not a definition. I suggest the following: "A set of graphic symbols of common origin, employed as the basic letters, characters or logograms in the alphabet, in the syllabary or in the logographic lexicon of one or more languages."

At the end of the definition there is a part of a sentence not belonging here.

script, source; script, target

Transformed should be used instead of "converted" as a consequence of the term *names transformation*.

sequence; sequence rules, alphabetic

Not only alphabetic, but also syllabic and logographic sequence exists. As to the problems in the Roman script the most important one: existence of different alphabets with different sequences, the need for toponymic indexes and gazetteers using exclusively the basic characters as sequence units. In my experience the sequence of the Czech and Slovak *ch* after *h*, and the Swedish *ä and å* at the end of the alphabet, are the extremest examples, not to speak about the Hungarian digraphs *cs, gy, ly, ny, sz, ty* and *zs*, to be ordered as separate units after *c, g, l, n, s, t, z* respectively.

syllable

The example Br-no could be helpful in showing that the speech, and not (always) the writing is a basis.

toponomastics

The definition must be a mistake. It is a synonym for *toponymy*. Cf. *onomastics*.

transformation, names

This categorization does not work in practice, as a great number of *exonyms* were produced by translation.

translation

It should be mentioned, that the *meaning* of a toponym can be explained in different ways. According to one aspect, eg. the meaning of a geographical name is not the one expressed by

the constituent words, but the *relation* between the feature and the name as a sign. On the other hand, the *translation* of a *geographical name* is always based on the other type: meaning of the constituent words.

transliteration

The definition is wrong. It is not the aim of *transliteration* to preserve phonological elements. In case of an *alphabet* like the English or French as *source scripts*, a transliteration cannot respect the sounds in order to make possible the complete reversibility, because the alphabet is used inconsistently in the *source script* in respect to *phonemes* of the language.

If *transliteration* "aims at complete reversibility" then "in principle" should be omitted from the first sentence. Consequently, the example Ščukno for Щукно (correctly: Щукино) is wrong, as beside the šč (щ), there are š (ш) and č (ч) *transcriptions* in this system, thus reversibility is not possible. In case of *transliteration* щ = ŝ, ю = û, я = â instead of the UN recommendation šč, ju, and ja respectively.

The example Adis Abeba would be correct as the doubling is not shown in the Amharic script, but the vowel e represents an /ɛ/ sound (1st) that should be underlined to distinct it from the /e/ sound (5th) according to UN recommendation. The example Piraievs would be correct if there is a system (I never heard about it) which represents the Greek letter υ consequently with the Roman v; we have to bear in mind that an approved UN recommendation exists for the *transcription* of Greek, where this Greek letter should be transcribed variously as u, f, and v depending on the represented *phoneme*.

Theere is a problem in the background: the fact that two different aspects of *transliteration* exist. *Transliteration* in the strict sense, represented by ISO transliteration standards, realizes complete *reversibility*. On the other hand, there is a cartographic aspect which does not want a complete *reversibility*, but a systematic substitution of the source characters by the target characters (one character or combination of characters in the target script for one character or combination of characters in the source script where a character or a combination of characters can be used only once in the target script.

The present definition does not or is intended not to decide on this matter. However, the decision is inevitable. It must be stated either that *transliteration* is not a requirement in UN recommendations, i.e. *transcriptions* can also be approved, or to or to create another definiton for *transliteration* which reflects our aspect, where *reversibility* is not mentioned, and this type

of *transliteration* is considered as one in the broad sense, as opposite to the ISO system in the strict sense.

typeface

As already mentioned, the term *font* is used for this definition on computer printers.

UTM grid

Why not *topographic coordinates*? Cf. there.